UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

FEDERAL HOUSING FINANCE AGENCY, AS CONSERVATOR FOR THE FEDERAL NATIONAL MORTGAGE ASSOCIATION AND THE FEDERAL HOME LOAN MORTGAGE CORPORATION, Plaintiff, -v- UBS AMERICAS, INC., *et al.* Defendants.	**Case No. 11 CIV. 5201 (DLC)** **PLAINTIFF FHFA'S PROPOSED SAMPLING PROTOCOL**

Other Cases Brought By Plaintiff:

11 Civ. 6188 (DLC)
11 Civ. 6189 (DLC)
11 Civ. 6190 (DLC)
11 Civ. 6192 (DLC)
11 Civ. 6193 (DLC)
11 Civ. 6195 (DLC)
11 Civ. 6196 (DLC)
11 Civ. 6198 (DLC)
11 Civ. 6200 (DLC)
11 Civ. 6201 (DLC)
11 Civ. 6202 (DLC)
11 Civ. 6203 (DLC)
11 Civ. 6739 (DLC)
11 Civ. 7010 (DLC)
11 Civ. 7048 (DLC)

Pursuant to this Court's Order dated February 6, 2012, in the above-captioned cases (the

"Cases"), Plaintiff Federal Housing Finance Agency ("FHFA" or "Plaintiff"), as Conservator for

the Federal National Mortgage Association ("Fannie Mae") and the Federal Home Loan

Mortgage Corporation ("Freddie Mac") (together, the "GSEs"), respectfully submits this report

and proposal, as directed by the Court, regarding the use of statistical sampling in the Cases.

I. SUMMARY

As set forth in the sections that follow, FHFA proposes the use of statistical sampling, consistent with accepted methodologies and scientific principles, for purposes of discovery and trial in the Cases. FHFA's proposal is based on the following principles:

- Statistical sampling is a well-established means of streamlining litigation involving voluminous evidence, and it is an essential tool for conducting discovery and trial here. The Cases involve over a million loan files, consisting of hundreds of millions of pages, across 449 securitizations. Sampling will insure that discovery is manageable in scope, and that trials proceed without unreasonable, indeed extraordinary, burdens on the Court and jury.

- FHFA will analyze samples of randomly selected loans from securitizations in each Case, and extrapolate the results of its analysis to establish the extent of Defendants' false statements, including in each of three categories: (1) supposed adherence to underwriting guidelines; (2) loan-to-value ("LTV") ratios; and (3) owner-occupancy rates for the mortgaged properties. FHFA will draw statistically valid random samples sufficient to establish a 95% confidence interval, with a margin of error of ±5%. The 95% confidence interval has been generally accepted by courts and those in the scientific community. And FHFA will, as warranted, select samples from clusters of securitizations with similar characteristics, and stratify samples to permit findings addressed to the particular loan features at issue in the Cases.

- FHFA's selection of appropriate clusters and criteria for stratification depends in part on: (1) data contained on final closing loan tapes Defendants provided to the Trustees for each securitization that describe the features of the loans in each securitization; and (2) information also in Defendants' possession identifying the originators of the loans. FHFA therefore respectfully requests that Defendants be directed to produce the tapes provided to the Trustees and information sufficient to identify the originator of each loan to permit FHFA to generate a more detailed proposal – one that will identify specific sample sizes, specific securitizations suitable for clustering, and specific criteria for stratification. The parties can then exchange specific proposed sampling protocols, on a schedule set by the Court, and pursue possible agreement on an appropriate protocol, or, if necessary, seek the Court's intervention.

II. THE NECESSITY OF RELYING ON STATISTICAL SAMPLING

As articulated in FHFA's Proposal for Certain Case Management Issues, dated January 10, 2012, statistical sampling is key to the efficient management of these Cases and, ultimately, the presentation of evidence at trial. The Cases involve 449 securitizations. *See* Exhibit A

2

(listing number of securitizations per case). Even under the conservative assumption that the relevant population of loans is limited to those in the Supporting Loan Groups (that is, the loan groups collateralizing the certificates purchased by the GSEs in those securitizations),[1] there are over a million loans. According to Defendants' own estimate, an average loan file is 300 pages. (Transcript of Dec. 2, 2011 Hearing ("Tr.") at 34:14-17.) Producing and reviewing, on a loan-by-loan basis, over three hundred million pages of documentation would be a massive undertaking over an extended period of time. Nor would it be practical to present evidence on over a million loans to the fact-finder at the trials of these Cases.

At the December 2 hearing, after Defendants suggested that they may deem as many as 2.7 million loans to be at issue,[2] this Court recognized that "it would be extraordinarily burdensome and inefficient to reassemble, gather, review and analyze all of these files. … [A]s a practical matter, the burden on the defendants and the plaintiffs of personally and individually reviewing 2.7 million loan files, each of which has at least 300 pages, says there has to be a better way." (Tr. 34:6-11, 35:25-36:1, 38:18-21.) Statistical sampling is that "better way" – a reliable, efficient, and scientifically sound way for the parties to conduct discovery and adduce proof as to the loans at issue without reviewing each and every loan file.

III.   STATISTICAL SAMPLING IS A GENERALLY ACCEPTED METHOD FOR CASE MANAGEMENT

Statistical sampling is a well-established and scientifically valid methodology. Sampling is founded in accepted mathematical theory and has been in use for more than 200 years. Courts in this and other districts have consistently held that, provided the methodology is reliable,

---

[1]   FHFA may determine to draw for its samples loans from other loan groups as well as the Supporting Loan Groups, including for the purpose of rebutting any claim that the larger loan groups are the relevant population.

[2]   Defendants' figure appears not to be limited to loans in the Supporting Loan Groups but instead includes the loans in all loan groups in the securitizations.

sampling may be used to prove many different elements of a claim, including both liability and damages. *See, e.g., In re World Trade Ctr. Disaster Site Litig.*, 598 F. Supp. 2d 498, 505 (S.D.N.Y. 2009) (*sua sponte* deciding to use sampling for both liability and damages in adjudicating over 9,000 claims brought under Air Transportation Safety and System Stabilization Act by workers who cleaned up site of the September 11 attacks); *Rosado v. Wyman*, 322 F. Supp. 1173, 1180 (E.D.N.Y. 1970) ("Sampling has long been considered an acceptable method of determining the characteristics of a large universe. ... Such mathematical and statistical methods are well recognized by the courts as reliable and acceptable in determining adjudicative facts."), *aff'd*, 437 F.2d 619 (2d Cir. 1970); *see also Maran Coal Corp. v. Societe Generale de Surveillance S.A.*, 1996 U.S. Dist. LEXIS 174 (S.D.N.Y. Jan. 9, 1996) (Cote, J.) (rejecting defendant's application for pretrial hearing concerning admissibility of expert testimony regarding samples of cargo taken pursuant to contract at loading and discharge ports).

In particular, courts have approved sampling of loans in cases involving mortgage-backed securities. *See MBIA v. Countrywide*, No. 602825/08, 30 Misc. 3d 1201(A) (N.Y. Sup. Ct. Dec. 22, 2010) (approving plaintiff's request to use sampling in mortgage-backed securities case to support its fraud and contract claims); *Syncora Guarantee Inc. v. EMC Mortg. Corp.*, 09 Civ. 3106 PAC, 2011 WL 1135007 (S.D.N.Y. Mar. 25, 2011) (ruling that plaintiff can use statistical sampling to prove claim that defendant breached contract representation that loans were originated in violation of underwriting guidelines).

Further, sampling is commonly used to determine characteristics or attributes of financial instruments, including loans and RMBS. Federal regulators often sample loans held by financial institutions to determine whether and to what extent an institution is in compliance with its statutory and regulatory requirements. For example, as the Federal Deposit Insurance

Corporation ("FDIC") describes in a manual published on its website, its examiners use statistical sampling of a bank's loan portfolio to determine, among other things: (1) the bank's adherence to its own lending policies; (2) the adequacy of the quality of the bank's assets and collateral; and (3) whether the bank has charged the right interest rate and set aside the proper reserves.[3] Private businesses likewise routinely rely on sampling of mortgage loan pools, for the purpose of assessing credit quality, compliance with regulatory requirements and laws, and adherence to internal policies and procedures. Loan originators routinely use statistical sampling to conduct internal quality control, compliance, and fraud audits, and to value the loans on their balance sheets. Finally, securities underwriters, including each of the underwriter Defendants in the Cases, relied on statistical sampling when conducting or commissioning due diligence of the loans in the securitizations.[4]

IV.   OBJECTIVES OF USING SAMPLING IN THE CASES

FHFA proposes to use statistical sampling to support its allegations that the Prospectus Supplements contained material misstatements and omissions concerning key characteristics of the underlying mortgage loans. Specifically, through the use of sampling, FHFA proposes to establish the falsity of Defendants' representations that: (1) loans were originated in compliance with the underwriting guidelines of the originators; (2) the properties collateralizing certain

---

[3]   *See* FDIC Risk Management Manual of Examination Policies, available at http://www.fdic.gov/regulations/safety/manual/toc_all.html.

[4]   *See, e.g.*, Banc of America Mortgage Securities Inc., Prospectus Supplement for Banc of America Alternative Loan Trust 2006-3 (SEC Form 424B5), at 31 (March 30, 2006) ("Generally, Bank of America conducts a post-purchase review of a sampling of all mortgage loans acquired from another lender to determine whether agreed upon requirements were met."); Credit Suisse First Boston Mortgage Acceptance Corp., Prospectus Supplement for CSFB Mortgage-Backed Trust Series 2005-11 (SEC Form 424B5), at 35 ("Additionally, all or a sample of the loans comprising mortgage assets for a series may be reviewed by or on behalf of the depositor to determine compliance with those underwriting standards and procedures and compliance with other requirements for inclusion in the trust.").

numbers of loans were owner-occupied; and (3) certain numbers of loans had LTV ratios of specified values. In particular, after representative samples are selected, FHFA will: (1) analyze each loan file in the sample to identify loans originated in violation of the originator's underwriting guidelines, false statements of owner-occupancy rates, and false statements of LTV ratios; and (2) extrapolate from each sample the incidence of such defects and false statements across the corresponding population. Sampling and extrapolation are also relevant to calculating damages on claims alleged in each of the Complaints.

Sampling is also relevant to proving fraud because, in addition to establishing the falsity of representations, the extent and correlation of misrepresentations and patterns in the incidence of defects can tend to prove scienter or intent. *See Arista Records LLC v. Usenet.com, Inc.*, 633 F. Supp. 2d 124, 152 (S.D.N.Y. 2009) ("staggering scale of infringement" supported finding of willfulness); *see also United States v. Conner*, 262 F. App'x 515 (4th Cir. 2008) (sampled claims probative of intent in healthcare criminal fraud case). To take the most basic example, a higher rate of defects and misrepresentations is more consistent with deliberate disregard of the underwriting guidelines than mistake or accident. Similarly, if the sample demonstrates that the distribution of underwriting guidelines violations is significantly higher for borrowers with lower FICO scores than for borrowers with higher FICO scores (that is, if the distribution of underwriting guideline violations is non-random and tilted towards lower FICO scores), that is a strong indication of fraud, because the pattern is consistent with originators deliberately violating guidelines for borrowers who otherwise would be less qualified to obtain the mortgage. Mistaken or unintentional failures to comply with the guidelines would be more consistently distributed.

V.   PROPOSED STATISTICAL SAMPLING METHODOLOGY IN THE CASES

FHFA intends to take the following three accepted and conventional steps in sample design: (1) choosing the desired confidence level and margin of error; (2) selecting the sample to generate that confidence level; and (3) extrapolating from the sample to the population. FHFA's proposals on these steps are as follows:

A.   Confidence Level

FHFA will select samples that achieve a confidence interval of 95 percent, with a margin of error of plus or minus 5 percent. A "confidence level" is a measure of the accuracy with which a sample can be used to estimate values in the population from which the sample is drawn. It refers to the percentage of times that the true value of the population will fall within a particular range based on results from the sample.

The proposed confidence interval of 95 percent is customary in statistics.[5] This confidence interval means that, if a sample of the same size were selected from the population a very large number of times, the true population value would be found inside the sample range defined by the margin of error 95 percent of the time. As an illustration, assume that 85 percent of the mortgage loans in a sample drawn from a securitization failed to comply with the originator's underwriting guidelines. If the sample size yields a confidence level of 95 percent with a margin of error of plus or minus 5 percent, the defect rate for the loans in the population would be between 80 and 90 percent, with only a 2.5 percent chance that the true value of defects in the population is below 80 percent and a 2.5 percent chance that it is above 90 percent.

---

[5] *See* Kevin D. Hoover & Mark V. Siegler, *Sound and Fury: McCloskey and Significance Testing in Economics,* 15 J. Econ. Methodology 1, at 13-14, 24 (March 2008) ("The critical value is typically but not always chosen to secure a 5% probability of type I error under the null hypothesis (i.e. a 5% size of the test)" and "...epidemiology or other areas of medical research ... faithfully apply a standard of $p < 0.05$ for reporting estimates").

In order to propose a specific sample size for use in any particular Case, it is necessary

first to define the population. To give an example: to achieve the 95 percent confidence interval

for a population of approximately 4,000 loans or greater, a sample of 384 (rounded up to 400 to

provide a modest margin) is sufficient under accepted mathematical principles. *See* Steven K.

Thompson, *Sampling* 42 (2d Ed. 2002) (*see* example 1). However, if the population is smaller, a

smaller sample size will suffice. Here, to know the size of the populations, the parties should

agree if possible on, among other things, whether the relevant populations will consist of only

those loans in the Supporting Loan Groups or, more broadly, all loan groups in a Securitization.

There are multiple ways to draw a statistically valid random sample of loans with a 95

percent confidence interval. Two available tools are stratification and clustering. FHFA will

best be able to determine appropriate criteria for employing these methods based on data in the

loan tapes that Defendants provided to the Trustees for each securitization[6] and, to the extent it is

not included on the loan tapes, information that identifies the originator of each loan.

B.   Stratification

FHFA intends to examine the characteristics of loans in the relevant population to

determine whether and how to stratify the population. For a population that is not stratified,

selection of the sample is random across the entire population. Stratification is the division of a

population into mutually exclusive subgroups, known as "strata," by using characteristics of that

population. For example, one could create subgroups based on FICO score ranges (620-674;

675-699; 700-719; and 720-850). From each of these subgroups, a proportionate random sample

---

[6]   The closing loan tapes delivered to the Trustees for the securitizations are generally the
last and most updated versions of loan tapes generated and used during the securitization
process. The loan tape typically takes the form of data file, such as an Excel spreadsheet, that
contains approximately 50-60 data fields regarding characteristics for each loan in the
Securitization.

is drawn. Thus, if one was drawing a sample of 20 loans from a population of 100 and there were four strata of 25 loans each, then a random sample of five loans would be drawn from each of those four groups to form the sample. Stratification is commonly used in sampling to attempt, among other things, to increase the precision of the estimates generated from the sample by reducing variability, and to enhance the ability to make separate estimates about subgroups within the population.

Upon examining data obtained from the loan tapes and the originator information, FHFA will consider stratifying each sample based on the following factors, among others: reported LTV/Combined LTV ("LTV"), loan performance, documentation type, occupancy type, and FICO score. Each factor relates to values FHFA intends to measure with the sample. For example, the measurement of the incidence of false LTV/CLTV data in the Prospectus Supplements is related to the reported LTV/CLTV data on the loan tapes.

A basic requirement for stratification is information about the characteristics of each loan in the population. One cannot stratify based on a particular factor unless that factor is known for each member of the population. Here, the information concerning the individual loans in the population is most readily available on the loan tapes provided to the Trustees, which contain statistics for each loan, such as FICO score, reported LTV/CLTV, and other data.

C.   Clustering

FHFA will also evaluate the possible use of clustering to achieve the desired confidence level. Clustering places homogeneous subpopulations into a single group, or cluster. In these Cases, the clustering would be of securitizations within a Case. A compelling reason to consider clustering is that even within a single Case the number of securitizations may render it infeasible to proceed by sampling each securitization separately. For example, *FHFA v. Merrill Lynch* concerns 72 securitizations. Sampling 400 loans from each of the 73 Supporting Loan Groups in

9

that Case will still obligate the parties to analyze, and the fact-finder to consider, over 29,000 loans. In *FHFA v. Deutsche Bank*, there are 40 securitizations, meaning that if 400 loans from each Securitization are sampled, the sample will comprise 16,000 loans.

Clusters could be organized if the loan groups of more than one securitization are sufficiently homogenous as to key factors. For example, FHFA will consider whether to propose clustering securitizations where the originator or major originators were the same and the loans were originated during the same timeframe.

Clustering allows the sample to be drawn from one or two randomly selected securitization(s) from each cluster. That sample would be used to generate estimates for the cluster. It is also possible to use stratification in conjunction with clustering, such that loans from each securitization are within the sample. That is, one can draw loans from each securitization in the cluster and stratify by securitization. Extrapolated values for the securitizations and for the cluster can then be generated, with the degree of precision being greater for the cluster than for the securitization.

FHFA requires additional information to determine whether, and how, to use clustering. Such an analysis requires a close study of the characteristics of the securitizations and their similarity to each other. FHFA can most readily assess these similarities with access to the loan tapes and information identifying the originators of the loans.

D.   Extrapolation

Once a sample of loans is selected, and the loans analyzed, the values obtained for the sample can be estimated for the population. This is known as extrapolation. The appropriate extrapolation technique is determined by the sample design and the results of the review of the sample, however, and thus cannot be specified without, among other things, additional information about the characteristics of the loans in the pools, including originator information.

VI.   TIMETABLE FOR CASE-SPECIFIC PROTOCOLS AND RELIEF SOUGHT

FHFA is prepared to propose specific sample selection criteria for the samples in each Case, shortly after it receives loan tapes and originator information for the securitizations, and will consider any counter-proposals from Defendants in the interests of achieving possible agreement between the parties. The loan tapes provided to the Trustees should be readily available, as they are a basic piece of documentation relating to any securitization. Originator information, on a loan-by-loan basis, should also be in the possession of Defendants. FHFA estimates that it would be able to present a specific sampling protocol for the first Case for which it receives the requested information—a protocol describing the precise parameters for sample size, stratification, and clustering—within five weeks of production of that information. It can generate samples for additional Cases, tailored as necessary to the loans at issue in those Cases, on a rolling basis thereafter.

FHFA also respectfully requests that Defendants be ordered to produce Case-specific protocols, samples, and methodologies, on the same timetable as FHFA. In the parties' numerous meet and confers, Defendants have articulated no principled objection to the use of sampling and instead have taken the position that it is premature to consider the issue. To defer the issue as Defendants prefer, however, is to predetermine that the parties will be compelled to analyze individual loan files for every disputed securitization, a monumental undertaking given the volume of loans at issue in each Case. The appropriate time to settle on an appropriate sampling protocol is now, at the outset of the Cases, and Defendants should participate in that process by proposing Case-specific protocols on the same timetable as FHFA.
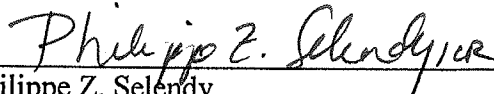
*   *   *

11

In light of the foregoing, FHFA respectfully requests that this Court enter an Order:

1) Requiring each Defendant who has possession, custody, and control, of loan tapes and loan-level originator information for the securitizations in each Case, to produce such information in the Case in which it is named by April 5, 2012[7]; and

2) Requiring that the parties simultaneously exchange the protocol for at least one Case before May 10, 2012, and confer about the schedule for simultaneously exchanging Case-specific sampling protocols in the remaining cases on a rolling basis, with those exchanges to be completed by June 14, 2012.
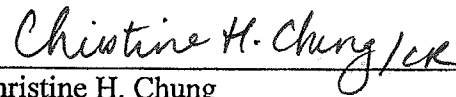
---

[7]   FHFA will coordinate with Defendants so as to not require that the same loan tape be produced multiple times.

Dated: February 29, 2012
New York, New York

Respectfully submitted,

_____
Philippe Z. Selendy
(philippeselendy@quinnemanuel.com)
QUINN EMANUEL URQUHART &
    SULLIVAN, LLP
51 Madison Avenue, 22nd Floor
New York, New York 10010
(212) 849-7000

*Attorneys for Plaintiff Federal Housing Finance
Agency in FHFA v. UBS Americas, Inc., FHFA
v. JPMorgan Chase & Co., FHFA v. Deutsche
Bank AG, FHFA v. Citigroup Inc., and FHFA v.
Goldman, Sachs & Co.*

_____
Marc E. Kasowitz (mkasowitz@kasowitz.com)
Hector Torres (htorres@kasowitz.com)
Christopher P. Johnson
(cjohnson@kasowitz.com)
Michael Hanin (mhanin@kasowitz.com)
Kanchana Wangkeo Leung
(kleung@kasowitz.com)
KASOWITZ, BENSON, TORRES &
    FRIEDMAN LLP
1633 Broadway
New York, New York 10019

*Attorneys for Plaintiff Federal Housing Finance
Agency in FHFA v. Ally Financial Inc., FHFA
v. General Electric Company, FHFA v. Morgan
Stanley, and FHFA v. SG Americas, Inc.*

_____
Christine H. Chung
(christinechung@quinnemanuel.com)
QUINN EMANUEL URQUHART &
    SULLIVAN, LLP
51 Madison Avenue, 22nd Floor
New York, New York 10010

*Attorneys for Plaintiff Federal Housing Finance
Agency in FHFA v. First Horizon National
Corp., FHFA v. Bank of America Corp., and
FHFA v. Credit Suisse Holdings (USA), Inc.*

_____
Manisha M. Sheth
(manishasheth@quinnemanuel.com)
QUINN EMANUEL URQUHART &
    SULLIVAN, LLP
51 Madison Avenue, 22nd Floor
New York, New York 10010

*Attorneys for Plaintiff Federal Housing Finance
Agency in FHFA v. UBS Americas, Inc., FHFA
v. JPMorgan Chase & Co., FHFA v. Barclays
Bank PLC, FHFA v. Citigroup Inc., and FHFA
v. Merrill Lynch & Co., Inc.*

_____
Richard A. Schirtzer
(richardschirtzer@quinnemanuel.com)
Adam M. Abensohn
(adamabensohn@quinnemanuel.com)
QUINN EMANUEL URQUHART &

13

SULLIVAN, LLP
51 Madison Avenue, 22nd Floor
New York, New York 10010

Attorneys for Plaintiff Federal Housing Finance
Agency in FHFA v. HSBC North America Holdings, Inc.
 and FHFA v. Nomura Holding America, Inc.

EXHIBIT A

Complaint	Number of Securitizations
No. 11 Civ. 5201 (UBS)	22
No. 11 Civ. 6188 (JP Morgan)	103
No. 11 Civ. 6189 (HSBC)	17
No. 11 Civ. 6190 (Barclays)	8
No. 11 Civ. 6192 (Deutsche Bank)	40
No. 11 Civ. 6193 (First Horizon)	5
No. 11 Civ. 6195 (Bank of America)	23
No. 11 Civ. 6196 (CitiGroup)	10
No. 11 Civ. 6198 (Goldman Sachs)	40
No. 11 Civ. 6200 (Credit Suisse)	43
No. 11 Civ. 6201 (Nomura)	7
No. 11 Civ. 6202 (Merrill Lynch)	72
No. 11 Civ. 6203 (SocGen)	3
No. 11 Civ. 6739 (Morgan Stanley)	33
No. 11 Civ. 7010 (Ally (GMAC))	21
No. 11 Civ. 7048 (General Electric)	2